

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:23 ; Search time 299.73 Seconds
(without alignments)
26.982 Million cell updates/sec

Title: US-09-331-631A-5_COPY_76_144
Perfect score: 381
Sequence: 1 NR0RDPQOQYEOCCKRCORR.....EEOQREDEKYEERKESGDN 69

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 segs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL_15:*
2: sp_Archaea:*
3: sp_Bacteria:*
4: sp_Fungi:*
5: sp_Human:*
6: sp_Invertebrate:*
7: sp_Mhc:*
8: sp_Organelle:*
9: sp_Phage:*
10: sp_Plant:*
11: sp_Protozoa:*
12: sp_Virus:*
13: sp_Vertebrate:*
14: sp_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	381	100.0	625	10 Q9SP13	Q9SP13 macadamia i
2	363	95.3	666	10 Q9SP15	Q9SP15 macadamia i
3	359	94.2	666	10 Q9SP14	Q9SP14 macadamia i
4	173	45.4	525	10 Q43358	Q43358 theobroma c
5	166	43.6	593	10 Q9SEW4	Q9SEW4 juglans reg
6	115	30.2	810	10 Q9ZWT3	Q9ZWT3 cucurbita m
7	114.5	30.1	1038	5 Q60983	Q60983 dictyostell
8	109	28.6	554	5 Q9VPS3	Q9VPS3 drosophila
9	109	28.6	1089	12 Q40947	Q40947 drosophila
10	105.5	27.7	385	5 Q17909	Q17909 caenorhabdi
11	102	26.8	1162	12 Q98148	Q98148 kaposi's sa
12	101	26.5	1737	5 Q9TW28	Q9TW28 dictyostell
13	100	26.2	838	5 Q9YOC9	Q9YOC9 dictyostell
14	100	26.2	953	4 Q14151	Q14151 homo sapien
15	100	26.2	1129	12 Q9OR71	Q9OR71 kaposi's sa
16	100	26.2	1590	5 P91152	P91152 caenorhabdi
17	99	25.9	1390	5 Q77033	Q77033 dictyostell
18	98.5	25.9	2701	4 Q9Y520	Q9Y520 homo sapien
19	97.5	25.6	1027	4 Q9NST7	Q9NST7 homo sapien

20	97.5	25.6	1105	4 Q95819	Q95819 homo sapien
21	97.5	25.6	1175	4 Q75172	Q75172 homo sapien
22	97.5	25.6	1233	11 P97820	P97820 mus musculu
23	97.5	25.6	1257	4 Q95033	Q95033 homo sapien
24	97.5	25.6	1700	5 Q9VVA9	Q9VVA9 drosophila
25	97	25.5	378	2 Q31328	Q31328 borrelia bu
26	97	25.5	378	2 Q08264	Q08264 borrelia bu
27	97	25.5	1268	4 Q9UKD8	Q9UKD8 homo sapien
28	97	25.5	1276	4 Q9UKE1	Q9UKE1 homo sapien
29	97	25.5	1323	4 Q9UKE0	Q9UKE0 homo sapien
30	97	25.5	1331	4 Q9UKE4	Q9UKE4 homo sapien
31	96	25.2	355	5 Q9UOV4	Q9UOV4 Leishmania
32	96	25.2	1297	4 Q9UKD9	Q9UKD9 homo sapien
33	96	25.2	1305	4 Q9UKE3	Q9UKE3 homo sapien
34	96	25.2	1333	4 Q60298	Q60298 homo sapien
35	96	25.2	1352	4 Q9UKE2	Q9UKE2 homo sapien
36	96	25.2	1360	4 Q9UKE5	Q9UKE5 homo sapien
37	95.5	25.1	375	2 Q87298	Q87298 borrelia bu
38	95.5	25.1	483	10 Q49227	Q49227 pisum sativ
39	95.5	25.1	1026	5 Q9VLJ3	Q9VLJ3 drosophila
40	95.5	25.1	1295	4 Q9P2R8	Q9P2R8 mus musculu
41	95.5	25.1	1300	11 Q9JMG2	Q9JMG2 mus musculu
42	95.5	25.1	1303	4 Q9PIX1	Q9PIX1 homo sapien
43	95.5	25.1	1308	11 Q9UM52	Q9UM52 mus musculu
44	94.5	24.8	876	11 Q9WU62	Q9WU62 mus musculu
45	94	24.7	233	4 Q9UFP2	Q9UFP2 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	625 AA.
Q9SP13	Q9SP13			
AC	Q9SP13:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VICILIN PRECURSOR (FRAGMENT).			
GN	AMP2.			
OS	Macadamia integrifolia (Macadamia nut).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.			
OX	NCBI_TaxID=60698;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NOT KERNEL;			
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;			
RT	"A family of antimicrobial peptides is produced by processing of a 7S			
RT	globulin protein in Macadamia integrifolia.";			
RL	Plant J. 0:0-0(1999).			
DR	EMBL: AF161885; AAD54246.1; -.			
DR	HSSP: P02853; 2PHL.			
DR	INTERPRO: IPR001113; -.			
DR	PFAM: PF00546; Seedstore_7s; 1.			
FT	NON_TER			
DR	NON_TER			
SO	SEQUENCE	625 AA;	73586 MW;	415808A89D370296 CRC64;

Query Match 100.0%; Score 381; DB 10; Length 625;
Best local similarity 100.0%; Pred. No. 3.1e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 NR0RDPQOQYEOCCKRCORRETEPRHMQTCQGCERYEKKKKKKRYEEOQREDEKX 60
DB	
QY	76 NR0RDPQOQYEOCCKRCORRETEPRHMQTCQGCERYEKKKKKKRYEEOQREDEKX 135
DB	
QY	61 EERKKEGDN 69
DB	
QY	136 EERKKEGDN 144

RESULT 2

```

Q9SPL5
ID Q9SPL5 PRELIMINARY: PRT: 666 AA.
AC Q9SPL5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antitoxicobal peptides is produced by processing of a 7S
globulin protein in Macadamia integrifolia kernels.";
RL Plant J. 0:0-0(1999).
DR EMBL: AF161883; AAD54244.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78217 MW; C752B884B2DF0224 CRC64;

```

```

Query Match          95.3%; Score 363; DB 10; Length 666;
Best Local Similarity 95.7%; Pred. No. 1.4e-25;
Matches 66; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NRORPQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 60
DB 117 NRORPQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 176
QY 61 EERKKEGDN 69
DB 177 EERKKEEDN 185

```

```

RESULT 3
Q9SPL4 PRELIMINARY: PRT: 666 AA.
ID Q9SPL4;
AC Q9SPL4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antitoxicobal peptides is produced by processing of a 7S
globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL: AF161884; AAD54245.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78243 MW; 0ECA22F8710F8A7B CRC64;

```

```

Query Match          94.2%; Score 359; DB 10; Length 666;
Best Local Similarity 94.2%; Pred. No. 3.2e-25;
Matches 65; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NRORPQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 60
DB 117 NRORPQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 176

```

```

QY 61 EERKKEGDN 69
DB 177 EERKKEEDN 185

```

```

RESULT 4
Q43358 PRELIMINARY: PRT: 525 AA.
ID Q43358;
AC Q43358;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN CSV.
OS Theobroma cacao (Cacao).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Theobroma.
OX NCBI_TaxID=3641;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAVES;
RA MEDLINE=92288309; PubMed=1600151;
RA McHenry L., Filtz P.J.;
RT "Comparison of the structure and nucleotide sequences of vicilin genes
of cocoa and cotton raise questions about vicilin evolution.";
RL Plant Mol. Biol. 18:1173-1176(1992).
DR EMBL: X62625; CAA44493.1; -.
DR EMBL: X62626; CAA44494.1; -.
DR HSSP: P02853; 2PHL.
DR MENDEL: 30919; Thecc; 1188; 30919.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PRODOM: PD081059; -. 1.
DR Signal.
KW SIGNAL.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 525 VICILIN.
SQ SEQUENCE 525 AA; 60798 MW; 19114CD5C248905D CRC64;

```

```

Query Match          45.4%; Score 173; DB 10; Length 525;
Best Local Similarity 32.6%; Pred. No. 1.6e-08;
Matches 31; Conservative 17; Mismatches 15; Indels 32; Gaps 1;

QY 3 QRPDQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 46
DB 35 QRPDQOQYEOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 94
QY 47 -----KRYEEOOREDEEYERMK 65
DB 95 QOQGOQOQOCCRCORRETEPRHMOICOCRCERYEKKRKKOORRYEEOOREDEEY 129

```

```

RESULT 5
Q9SEW4 PRELIMINARY: PRT: 593 AA.
ID Q9SEW4;
AC Q9SEW4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS Juglans regia (English walnut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Juglandaceae; Juglans.
OX NCBI_TaxID=51240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SUNLAND; TISSUE=SONATIC EMBRYO LINE;
RA Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT "Identification and cloning of a cDNA encoding a vicilin-like protein,
Jug r 2, from English walnut kernel (Juglans regia): a major food

```


RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003588; AAF51469.1; -
 DR FLYBASE: FBgn031273; CG2839.
 DR INTERPRO: IPR001304; -
 DR PIRAM: PF00039; lectin_c: 1.
 DR PROSITE: PSS0041; C_type_LectIN_2; 1.
 FT NON-TER 554 554
 SQ SEQUENCE 554 AA; 69873 MW; 029BD852609FD16 CRC64;

Query Match 28.6%; Score 109; DB 5; Length 554;
 Best Local Similarity 32.3%; Pred. No. 0.01;
 Matches 21; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

OY 2 RORDPOOQYEOCCKRCORRETEPRHMQICQRCERYEKKRKQKRYEEOQREDEKYE 61
 DB 378 KKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKE 437
 OY 62 ERKKE 66
 DB 438 ERRE 442

RESULT 9
 O40947 PRELIMINARY; PRT; 1089 AA.
 AC O40947;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE ORF 73.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=37296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97296220; PubMed=9151804;
 RA Neipel F., Albrecht J.C., Fleckenstein B.;
 RT "Cell-homologous genes in the Kaposi's sarcoma-associated thadnovirus
 RT human herpesvirus 8: determinants of its pathogenicity?";
 RL J. Virol. 71:4187-4192(1997).
 DR EMBL: U93872; AAB62657.1; -
 DR INTERPRO: IPR002017; -
 SQ SEQUENCE 1089 AA; 126232 MW; 48AE2F9837E1AB6A CRC64;

Query Match 28.6%; Score 109; DB 12; Length 1089;
 Best Local Similarity 32.3%; Pred. No. 0.018;
 Matches 21; Conservative 24; Mismatches 20; Indels 0; Gaps 0;
 OY 2 RORDPOOQYEOCCKRCORRETEPRHMQICQRCERYEKKRKQKRYEEOQREDEKYE 61
 DB 378 KKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKEERKKKE 437

DB 589 QOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQO 648
 OY 62 ERKKE 66
 DB 649 QELE 653

RESULT 10
 O17909 PRELIMINARY; PRT; 385 AA.
 AC O17909;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE C1166.3 PROTEIN.
 GN C1166.3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkinson J.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Coop-r J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shonkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaadin M., Vaughan K., Waterson R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";
 RL Nature 368:32-38(1994).
 DR EMBL: Z70204; CAA94113.1; -
 DR INTERPRO: IPR001965; -
 DR PIRAM: PF00628; PHD; 1.
 SQ SEQUENCE 385 AA; 46052 MW; D080FC63B47EDC1 CRC64;

Query Match 27.7%; Score 105.5; DB 5; Length 385;
 Best Local Similarity 28.4%; Pred. No. 0.016;
 Matches 21; Conservative 24; Mismatches 20; Indels 9; Gaps 1;
 OY 2 RORDPOOQYEOCCKRCORRETEPRHMQICQRCERYEKKRKQKRYEEOQREDEKYE 52
 DB 146 RERDEKEREQKQKEREKRDARRELEKAKENDAKRYAVEEERKKKEKKKEKKKE 205
 OY 53 OREDEKYEERKKE 66
 DB 206 LTKERKESERKEKE 219

RESULT 11
 O98148 PRELIMINARY; PRT; 1162 AA.
 AC O98148;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE ORF73 HOMOLOG.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=37296;
 RN [1]
 RP SEQUENCE FROM N.A.

